

## 拒絶理由通知書

特許出願の番号	平成11年 特許願 第154308号
起案日	平成15年 4月28日
特許庁審査官	小林 紀和 4240 5X00
特許出願人代理人	高橋 詔男 (外 3名) 様
適用条文	第29条第2項

この出願は、次の理由によって拒絶をすべきものである。これについて意見があれば、この通知書の発送の日から60日以内に意見書を提出して下さい。

## 理 由

A. この出願の請求項 1-22 に係る発明は、その出願前日本国内又は外国において頒布された下記 1 及び 2 の刊行物に記載された発明に基いて、その出願前にその発明の属する技術の分野における通常の知識を有する者が容易に発明をすることができたものであるから、特許法第29条第2項の規定により特許を受けることができない。

## 記

1. 特開平10-336249号公報 (段落17、図3を参照)
2. 特開昭63-197148号公報 (単パケットから長パケットを作成)

## 先行技術文献調査結果の記録

- ・ 調査した技術分野  
国際特許分類第7版 (IPC 7): H04L 12/  
Fタームテーマ : 5K030 (広域データ交換)
- ・ 先行技術文献  
特開平02-195756号公報

この先行技術文献調査の記録は、拒絶理由を構成するものではない。

この拒絶理由通知書の内容等に関する問い合わせ先

特許審査第四部 デジタル通信 (データネットワーク) 小林紀和

発送番号 1 4 7 9 2 1

発送日 平成 1 5 年 5 月 6 日 2 / 2

---

電話 (03) 3581-1101 内線3556

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-336249  
 (43)Date of publication of application : 18.12.1998

(51)Int.Cl. H04L 12/56  
 H04L 12/28  
 H04Q 3/00

(21)Application number : 10-137838 (71)Applicant : AT & T CORP  
 (22)Date of filing : 20.05.1998 (72)Inventor : CIVANLAR SEYHAN  
 SAKSENA VIKRAM R

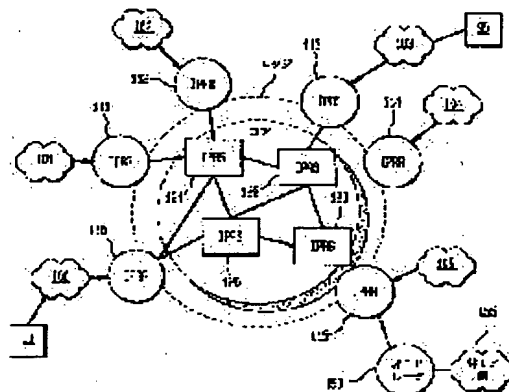
(30)Priority  
 Priority 97 859315 Priority 20.05.1997 Priority US

## (54) INTERNET PROTOCOL REPEATING NETWORK

## (57)Abstract:

PROBLEM TO BE SOLVED: To reduce the number of router hops and the amount of layer 3 processing in each hop by transferring an internet protocol(IP) packet based on destination information contained in a label added to the packet.

SOLUTION: The IP packet is sent from a host 90 through a LAN 100 to an entrance IP repetition route (IPRR) 110. The IPRR 110 executes layer 3 processing concerning the packet and at the same time, the label, which shows the router of the next hop is an exit IPRR 113, is added to the packet. The packet is transferred to an IP repeating switch (IPRS) 120. The IPRS 120 reads the added label and transfers the packet to an IPRS 122. The IPRS 122 also reads this label and transfers the packet to an exit IPRR 113. The exit IPRR 113 receives the packet, reassembles the packet by executing layer 3 processing and transfers this packet through a LAN 103 to a destination host 95.



## LEGAL STATUS

[Date of request for examination] 12.03.1999

[Date of sending the examiner's decision of rejection] 16.08.2000

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection] 2000-18089

[Date of requesting appeal against examiner's decision of rejection] 14.11.2000

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19)日本国特許庁 (J P)

(12) 公 開 特 許 公 報 (A)

(11)特許出願公開番号

特開平10-336249

(43)公開日 平成10年(1998)12月18日

(51)Int.Cl.<sup>8</sup>

識別記号

F I

H 0 4 L 12/56

H 0 4 L 11/20

1 0 2 A

12/28

H 0 4 Q 3/00

H 0 4 Q 3/00

H 0 4 L 11/20

G

審査請求 未請求 請求項の数17 O L (全 12 頁)

(21)出願番号 特願平10-137838

(22)出願日 平成10年(1998)5月20日

(31)優先権主張番号 08/859315

(32)優先日 1997年5月20日

(33)優先権主張国 米国 (U S)

(71)出願人 390035493

エイ・ティ・アンド・ティ・コーポレーション

AT&T CORP.

アメリカ合衆国 10013-2412 ニューヨーク  
ニューヨーク アヴェニュー オブ  
ジ アメリカズ 32

(72)発明者 セイハン シヴァンラー

アメリカ合衆国 07701 ニュージャージー,  
レッド バンク, コールマン アヴェ  
ニュー 85

(74)代理人 弁理士 岡部 正夫 (外3名)

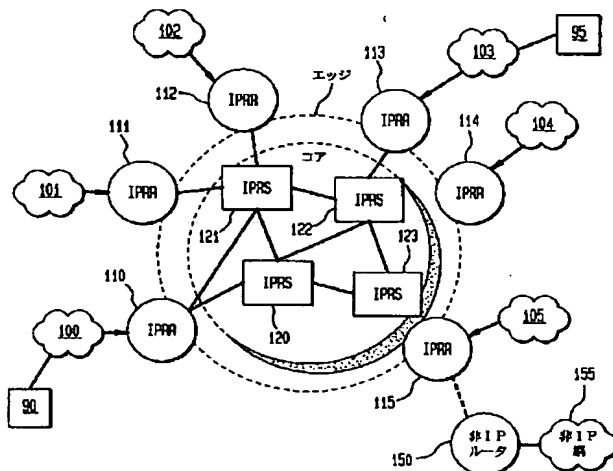
最終頁に続く

(54)【発明の名称】 インターネットプロトコル中継網

(57)【要約】

【課題】 データグラムの伝送路に沿って経由されるルータホップ数を低減すると同時に、各ホップにおける層3処理の量も低減する必要性が存在する。

【解決手段】 エッジとコアを含むIPトラヒックを伝送するためのインターネット中継システムおよび方法が開示される。エッジは、入ロルータおよび出ロルータを含み、コアは、スイッチ網を含む。入ロルータは、ソース網と通信し、ソース網からIPパケットを受信する。入ロルータは、各IPパケットにグローバル的に一意なラベルを付加し、このラベルが、IPパケットを網を横断して転送するのに用いられる。中継スイッチ網は、入ロルータと通信し、入ロルータからIPパケットを受信し、IPパケットをその伝送路に沿って、付加されたラベル内に含まれる宛先情報に基づいて転送する。出ロルータは、中継スイッチ網からIPパケットを受信し、これを宛先網に転送する。宛先網は、IPパケットを受信すると、これを意図される宛先に転送する。



## 【特許請求の範囲】

【請求項1】 インターネットプロトコルトラヒックを網内の伝送路に沿ってソースから宛先に伝送するためのインターネット中継システムであって、このシステムが：前記ソースと通信し、前記ソースから前記データグラムを受信するための入ルーター；および前記データグラムを前記網を横断して転送するために用いられる前記宛先をグローバル的に一意に識別するラベルを含み、このラベルが前記入ルーターの所で前記データグラムに付加され；このシステムがさらに前記入ルーターと通信し、前記データグラムを前記入ルーターから前記ラベル内に含まれる情報に基づいて受信するための中継スイッチ網を含み、このスイッチ網が前記データグラムを前記ラベル内に含まれる情報に基づいて前記伝送路に沿って転送し；このシステムがさらに前記データグラムを前記中継スイッチ網から受信し、前記データグラムを前記宛先に前記ラベル内に含まれる情報に基づいて転送するための出口ルーターを含むことを特徴とするインターネット中継システム。

【請求項2】 前記中継スイッチ網が複数の中継スイッチを含み、前記中継スイッチ網が前記データグラムを前記入ルーターから前記出口ルーターに前記出口ルーターが前記入ルーターからあたかも1ルータホップしか離れてないかのように転送することを特徴とする請求項1のインターネット中継システム。

【請求項3】 前記データグラムに付加される前記ラベルが各スイッチによって読み出し専用として処理されることを特徴とする請求項2のインターネット中継システム。

【請求項4】 前記各中継スイッチがメモリデバイス内に格納された転送テーブルを含み、この転送テーブルが前記データグラムの前記宛先に対応するルーティング情報を含み、前記各中継スイッチが前記ラベルを読み出し、前記ラベル内に含まれる前記宛先情報を前記転送テーブル内に含まれる前記ルーティング情報とマッチングし（照合し）、前記データグラムを前記伝送路に沿って前記出口ルーターに、前記ルーティング情報に基づいて転送することを特徴とする請求項2のインターネット中継システム。

【請求項5】 前記入ルーターがメモリ内に格納されたルーティングテーブルを含み、このルーティングテーブルが、前記宛先情報および前記出口ルーターに対応する識別情報を含むことを特徴とする請求項1のインターネット中継システム。

【請求項6】 前記中継スイッチ網内の前記複数のスイッチの内の前記データグラム伝送路に沿う前記出口ルーターと通信する一つが出口スイッチであって、前記入ルーターがメモリデバイス内に格納された前記出口スイッチに対応する識別情報を含むルーティングテーブルを含むことを特徴とする請求項2のインターネット中継システム。

ム。

【請求項7】 前記入ルーターがメモリデバイス内に格納された前記出口ルーターに対応する識別情報を含む転送テーブルを含むことを特徴とする請求項2のインターネット中継システム。

【請求項8】 前記入ルーターがメモリデバイス内に格納された前記出口スイッチ上の前記出口ルーターと通信するポートに対応する識別情報を含む転送テーブルを含むことを特徴とする請求項2のインターネット中継システム。

【請求項9】 前記入ルーターがメモリデバイス内に格納された前記出口スイッチに対応する識別情報を含む転送テーブルを含むことを特徴とする請求項2のインターネット中継システム。

【請求項10】 インターネットプロトコル（IP）パケットを中継網を横断して伝送するための方法であって、前記パケットがソースから発信され、宛先に転送され、この方法が：前記IPパケットを前記ソースから受信するステップ；前記IPパケットを前記ソースから入ルーターに前記IPパケット内に含まれるアドレス情報に基づいて転送するステップ；前記入ルーターの所で前記IPパケットをパージング（構文解析）することで前記宛先のアドレスを読み出すステップ；前記入ルーターの所で前記IPパケットに前記宛先のアドレスに関する識別情報を含むラベルを付加するステップ；前記入ルーターから前記IPパケットを交換網に転送するステップ；および前記交換網から前記IPパケットを出口ルーターに、前記出口ルーターが前記入ルーターからあたかも1ルータホップしか離れてないかのように転送するステップを含むことを特徴とする方法。

【請求項11】 さらに、前記出口ルーターの所で前記ラベルを前記IPパケットから除去するステップを含むことを特徴とする請求項10の方法。

【請求項12】 さらに、前記出口ルーターの所で前記IPパケットをリアセンブリングするステップを含むことを特徴とする請求項11の方法。

【請求項13】 さらに、前記出口ルーターから前記IPパケットを前記宛先に転送するステップを含むことを特徴とする請求項10の方法。

【請求項14】 前記入ルーターがメモリデバイス内にルーティングテーブルを含み、このルーティングテーブルが前記出口ルーターに対応するアドレス情報を含むことを特徴とする請求項10の方法。

【請求項15】 前記入ルーターがメモリデバイス内にルーティングテーブルを含み、このルーティングテーブルが前記宛先アドレスに対応する前記スイッチ網のアドレス情報を含むことを特徴とする請求項10の方法。

【請求項16】 前記ルーティングテーブルがさらに前記スイッチ網の前記出口ルーターと通信するポートに関する識別情報を含むことを特徴とする請求項15の方法。

## 3

【請求項17】 前記入ルータがメモリデバイス内に格納された転送テーブルを含み、この転送テーブルが前記宛先アドレスに対応する前記情報に関するアドレス情報を含み、この方法がさらに、前記宛先アドレスを、前記出ルータに対応する前記アドレス情報とマッピング（照合）するステップを含むことを特徴とする請求項10の方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、一般的には、通信システム、より詳細には、インターネットプロトコルトラフィックを現存の方法よりも、より高速および／あるいはより効率的に伝送することを可能にするインターネットプロトコル中継網に関する。

## 【0002】

【従来の技術】データ通信網、例えば、インターネットワーク（“インターネット”）においては、情報が、インターネットプロトコル（“IP”）データグラムあるいはパケットの形式にて、ソースから宛先に送られる。これらデータグラムは、インターネット上を“IPバックボーン”を経由して送られる。このIPバックボーンは、高速・高性能網であって、他の網を互いにリンクする。データグラムのソースと宛先は、例えば、異なる二人のユーザによって所有されるパーソナルコンピュータであり得る。典型的には、データグラムは、ソースから意図される宛先に向う伝送路上の多数の異なるルータを経由して送られる。ルータは、単なるホストコンピュータであって、網トラフィックをそれに接続された一つの網から別の網に、ソース網と宛先網が同一タイプであるか否かに関係なく転送する。ルータは、データグラムを受信すると、宛先ホストのアドレスを読み出し、データグラムの次の転送先のルータを決定する。ルータは、次に、データグラムのアドレス見出し部分内の現在のルータのアドレスを、次のルータのアドレスと置換した後に、データグラムをその伝送路に沿って送る。これは、一般に、“層3”処理として知られており、OSI（Open System Interconnection;開放型システム間相互接続）7層モデル、あるいはTCP/IP（Transmission Control Protocol/Internet Protocol;伝送制御プロトコル/インターネットプロトコル）モデルのいずれかのネットワーク層に関する。このルータからルータへの“ホップ”は、“ホップ・バイ・ホップ”プロセスとしても知られているが、このホッププロセスが、データグラムが宛先に接続されたルータに到着し、データグラムの配達完了するまで継続される。IPデータグラムは、インターネットを横断して宛先に到達するまでに、多数のルータを経由し、これらの各ルータがデータグラムについて上述の層3処理を遂行する。この処理スキームを、以降、“従来のIP”システムあるいはモデルと呼ぶものとする。この従来のシステムでは、短所とし

## 4

て、上述のホップ・バイ・ホップ伝送プロセスと、このプロセスと関連する各ルータにおける層3処理に起因して、IPパケットを網を横断して伝送する際に、大きな、しかも、予測不能な遅延が発生する。

【0003】“Classical IP Over ATM（従来のATMを通じてのIPの伝送）”として知られているもう一つの網プロトコルシステムは、ATM（非同期転送モード）網を通じてのIPトラフィックをサポートするための最も良く知られている処理モデルであり、最近IETF（Internet Engineering Task Force;インターネット技術に関する専門班）によって定義されたものである。図1は、このタイプの網を簡略的に示す。ここでは、一例として、ソースホスト10が、パケットを、ソースホスト10のローカルIPサブネット（“LIS”）の外側に位置する宛先ホスト20に送るものと仮定する。最初に、ソースホスト10は、パケットを、図1において経路1として示されるように、ATMスイッチ12を経由して、ルータ11に送る。ATM網を通じての経路1の設定は、ソースホスト10によって、ルータ11のIPアドレスをATMアドレスに翻訳することによって達成され、こうして、ルータ11への仮想経路が設定される。この従来のIP Over ATM（ATMを通じてのIP伝送）モデルは、APR（アドレス解決プロトコル）サーバ13を用意する。APRサーバ13は、ローカルIPサブネット（LIS）内に登録されている全てのホストおよびルータに対して、IPからATMアドレスへの翻訳を遂行する。この例においては、ソース10は、アドレス解決リクエストを、APRサーバ13に向けて、同一のローカルIPサブネット（LIS）内のルータ11のIPアドレスに対応するATMアドレスを得るために送る。これに回答して、APRサーバ13は、ルータ11に対応するATMアドレスに関する情報を含むアドレス解決応答を送り返す。その後、ソース10は、ATM網を横断して、ルータ11のATMアドレスに向けて経路1を設定し、ルータ11にパケットを送る。ルータ11は、パケット伝送における次のホップ先が、ルータ14であることを決定をし、APRサーバ16を用いてルータ14のATMアドレスを見つけ、パケットを、経路2として示されるように、ATMスイッチ12およびATMスイッチ15を経由して、ルータ14に送る。ルータ14は、同様に、APRサーバ19を用いて、ルータ17のATMアドレスを見つけ、パケットを、図1に経路3として示されるように、ATMスイッチ15およびATMスイッチ18を経由して、ルータ17に送る。ルータ17は、これに回答して、自身が宛先20への経路上の最終ホップに当たるルータである決定を行なう。ルータ17は、APRサーバ21を用いて宛先20のATMアドレスを見つけ、パケットを、宛先20に、図1に経路4として示されるように、ATMスイッチ18を経由して送る。

【0004】このタイプの網は、ATM網の上部においては従来のIPスイッチングを維持するが、ただし、下側のATM網は、単に、もう一つのOSI層2技術であると見做し、このために、ATMの長所を完全に活用していない。加えて、IPルータ11、14、17およびATMスイッチ12、15、18は、次のホップのアドレスを決定するために別個のプロトコルを実行し、このために、このタイプの網構成は効率が悪い。さらに、パケットの伝送路に沿う各ルータの全てが、パケットの宛先に向っての次のホップに当たるルータの決定を遂行する。このために、各ルータが、パケットについて層3処理を遂行することでIPの宛先アドレスを調べることと；ルーティングプロトコルによって決定され、各ルータ内に格納されているルーティングテーブルから、次のホップに当たるルータを決定することを要求される。さらに、このシステムの場合も、IPからATMアドレスを得るためのマッピングステップが、各ARPサーバ内で1度ずつ必要とされると共に、ソースホスト10から宛先ホスト20までの経路に沿って、ATMスイッチ仮想経路(SVC)を設定するステップが4度も必要とされ、このために、ルーティングスキームに、より多くのステップが必要となる。

【0005】その他の多数の網プロトコルソフトウェアスキームによって、インターネット伝送の高速化および高効率化が試みられている。これらシステムは、大きく2つのグループに分類することが可能であり、第一のグループのシステムは、中間ルータにおける層3処理の量を低減することを試み；第二のグループのシステムは、データグラム経路内のルータのホップ数を低減あるいは完全に排除することを試みる。以下に、この2つのタイプのシステムの概要を各グループについて示す。

【0006】“従来のIPモデル-近道ルーティング(Classical IP Model-Cut Through Routing)”として知られる一つの網処理スキームは、ルータホップを完全に排除することを試みる。ただし、このシステムは、多量のIPデータが、一つの特定のエンドポイントから別のエンドポイントに向けて伝送されるような状況においてのみ利用可能である。このタイプの処理システムが図2に簡略的に示される。このシステムでは、例えば、ソースホスト30が、宛先ホスト40のIPアドレスに対応するATMアドレスを決定することで、ソースホスト30から宛先ホスト40に向けて近道経路、すなわち仮想回路を設定し、これによって、ルータ41、42、43を回避する。データグラムに対する近道経路は、ソースホスト30から、ATMスイッチ31、ATMスイッチ32、およびATMスイッチ33を経由して宛先ホスト40に向うように定義される。

【0007】ただし、この近道経路をソース30から宛先40に向けて設定するためには、ソース30はパケットを従来のホップ・バイ・ホップ網(Classical Hop-by

-Hop網)を経由して送ることによって、宛先40のアドレスを取得することを要求される。現在、IETF(インターネット技術に関する専門班)は、NHRP(Next Hop Routing Protocol;次ホップルーティングプロトコル)と呼ばれる新たなアドレス解決プロトコルについて研究中である。このプロトコルは、IPからATMアドレスへのローカルIPサブネット(LIS)を横断してのマッピングを許し、近道経路をATM網を経由して設定することを可能にする。このNHRPモデルにおいては、上述の従来のホップ・バイ・ホップ網において用いられるARP(アドレス解決プロトコル)サーバと類似するNHS(Next Hop Server;次ホップサーバ)34、35、36が、一つあるいは複数のローカルIPサブネット(LIS)毎に採用される。ARPサーバ(アドレス解決プロトコルサーバ)とNHS(次ホップサーバ)との大きな差異は、ARPサーバは、ローカルIPサブネット(LIS)内の自身のクライアントと通信するのみであるが、NHSは、自身のホストに加えて他の複数のローカルIPサブネット(LIS)をサポートする複数の近隣NHSとも通信する点である。

【0008】この構成においては、ATMスイッチ31に接続されたソースホスト30は、自身のIP-ATMアドレスに、従来のホップ・バイ・ホップ網において用いられるARPサーバ(アドレス解決プロトコルサーバ)と類似する自身に指定されたNHS(次ホップサーバ)34を登録することで、NHRP(次ホップルーティングプロトコル)クライアントとなる。NHS34は、ルータ41あるいはATMスイッチ31内に格納される。NHRPクライアント、ここではソースホスト30が、データトラヒックを、省略時(ホップ・バイ・ホップ)において用いられる経路から近道経路に切り替えることを希望する場合、ソースホスト30は、最初に、宛先ホスト40のアドレスを、リクエスト/応答プロセスを用いて取得することが必要である。このプロセスは、ソースホスト30が、宛先40のIPアドレスの対応するATMアドレスへのマッピングを得るために、NHRPリクエストを、ルータ41内に格納されたNHS34に、経路1aを介して送ることによって開始される。NHS34は、このリクエストを受信する。ここで、NHS34は、宛先40が同一のLIS内にある場合、あるいは、前のアドレス解決リクエスト/応答プロセスによってそのアドレスマッピングを学習しており、宛先ホスト40のアドレスがキャッシュメモリ内に格納されている場合は、このアドレスマッピングをデータベース内に持つ。ただし、NHS34が宛先40のアドレスを持たない場合は、NHS34は、このNHRPリクエストを、経路2aを介して、ルータ42内に格納されたNHS35に転送することが必要となる。このプロセスが継続され、最終的には、NHRPリクエストが、経路3aを介して、ルータ43内に格納されたNHS36

に転送される。ここでは、NHS 36は、宛先40への伝送路上の最後のNHSであり、経路4aを介して、宛先40に到達する。宛先40のATMアドレスを運ぶNHRP応答は、ソース30に戻る逆の経路に沿って流れ、NHS 34によってソースホスト30に向けて転送される。その後、ソースホスト30は、このATMアドレスを用いて、宛先への仮想経路を、ソース30と宛先40との間の省略時の伝送路に沿うルータ41、42、43を経由することなく、直接に設定する。このように、ソース30と宛先40との間の近道経路に対するアドレスマッピングを得るためには、実質的に上述のホップ・バイ・ホッププロセスと類似するリクエスト／応答プロセスを、各ルータにおいて層3処理を用いて少なくとも一度は遂行することが要求される。

【0009】NHRPプロセスを用いた場合、経路がいったん設定されると、ルータホッピングおよび関連する層3パケット処理を回避することが可能であるが、ただし、ATMスイッチ網を通じて設定可能な近道仮想回路の数は限られている。この短所は、仮想回路破裂（Virtual Circuit Explosion）として知られており、多数のIPトラヒックのためにATMスイッチ網を通じて同時に設定される仮想回路の数が管理不能に大きくなった場合に発生する。さらに、典型的にはデータグラムは小さなために、ソース30と宛先40の間に近道経路を設定するために要求されるこのリクエスト／応答プロセスによって、通常は、ソースと宛先との間のその特定の通信シーケンスに対する伝送要件が満足され、これらの間の近道経路はもはや不要となる。さらに、NHRPは、ATM網にしか適用できず、標準のIPパケットサイズより、むしろ、より大きなデータの伝送に適する。

【0010】IPナビゲータ（IP Navigator）として知られているもう一つのソフトウェアルーティングスキームは、広域エリア網（WAN）スイッチングとIPルーティングを結合することで、網上を伝送される各IPパケットに対して、エンド宛先スイッチ、つまり、出口ルータと通信するスイッチ（の識別）を格納する。こうすることにより、網内のスイッチがIPルータとして機能する。いったんエンド宛先スイッチが決定されると、IPパケットが、事前に設定された“マルチポイント・ツウ・ポイントトンネル”を通じて、そのスイッチに転送される。このマルチポイント・ツウ・ポイントトンネル法においては、いったん仮想回路、すなわち近道が設定されると、アドレス情報が網内の他の全てのノードにブロードキャストされる。ただし、この方法においても、エンド宛先スイッチにおいて、IPパケットを転送するための適当なエンド宛先スイッチの出口ポートを決定するために、ルーティングテーブルの検索が遂行される。このソフトウェアモデルも、近道ルーティングスキームと同様に、ATM網にしか適用できず、仮想回路破裂問題も持つ。

# 【0011】

【発明が解決しようとする課題】これら従来の近道ルーティングを用いるIPモデル（Classical IP with Cut Through Routing）およびIPナビゲータ（IP Navigator）スキームは、IPパケットを網を横断して伝送する際に経由するルータホップの数を減少、あるいは、完全に排除することを目指す。ただし、これら網では、依然として、かなりの量の層3処理が要求される。従って、データグラムの伝送路に沿って経由されるルータホップの数を低減すると同時に、各ホップにおける層3処理の量も低減する必要性が存在する。

【0012】中間ルータにおける層3処理の量を低減する試みとしては、タグスイッチング法（Tag-switching）およびIPスイッチング法（IP-switching）の二例の網スキームが存在する。ただし、これら網スキームは、いずれもIPデータグラムの経路内のルータホップの数を低減することはない。タグスイッチング法は、ETFによって提唱される方法で、OSIモデルの層2（データリンク層）と層3（ネットワーク層）との間に、いわゆる“タグ（tag）”、すなわちシム層を挿入する。タグスイッチング法は、データリンク層がATMである場合は、53バイトのATMセル見出し部分の仮想経路識別子（VPI）と仮想回路識別子（VCI）のある部分をタグとして用いることを提唱する。ソースと宛先の間の各IPルータは、挿入されたタグを調べ、ローカル検索テーブルに基づいて、IPデータグラムを、どのポートに転送するかを決定する。各ルータは、自身のデータベース内にテーブル項目を設定し、この項目は、特定のタグに対して、第一のルータから発信されるデータグラムが、第二のルータに向うどの特定のポートに転送されるべきかを示す。加えて、第一のルータは、呼設定機能を遂行する。すなわち、第一のルータは、経路上の第二のルータに対して、“タグ配布プロトコル（tag distribution protocol）”を用いて、第二のルータに転送されようとしている入りIPトラヒックが、特定のタグ識別子を持つことを知らせる。このプロセスが各ルータにおいて次々と継続される。つまり、各ルータは、データグラムが宛先に到達するまで、タグを調べ、データグラムを転送すべき下流のルータを決定し、現存のタグを下流のルータのアドレスを示す別のタグと置換し、データグラムをそのアドレスに向けて転送する動作を継続する。この場合も、データリンク層がATMである場合は、各ルータは、パケットの見出し部分のVPI/VCI値を調べることで、転送先の決定を行なう。つまり、タグスイッチング法は、IPパケット伝送路上の各ルータの入りポートと出ポートの間で、VPI/VCI値を置換（スワッピング）する動作を伴う。また、タグスイッチング法は、上述のように、IPデータグラムについて要求される層3処理の量を低減することを目指す。ただし、伝送路上のルータホップの数につ



いては改善はみられない。

【0013】IPスイッチング法として知られているもう一つのシステムも、IPパケットに関して伝送の際に遂行される層3処理の量を低減することを目指す。この方法は、ATM網においのみ用いられる。IPスイッチング法の一般概念は、上述のタグスイッチング法と類似する。ただし、このIPスイッチング法は、タグスイッチング法において用いられるATMシグナリングおよび呼設定プロセスは用いない。ATMスイッチは、単に、IPデータグラムを、53オクテットのかたまりの形式で中継するために用いられ、この際に、各パケットに、短かなタグが付加される。このタグは様々なVPI/VCI値を含み、これらの使用は各ルータによってローカルの決定される。タグスイッチング法およびIPスイッチング法は、両方とも、タグを操作することで、各IPデータグラムに関する処理量を低減するが、ただし、ルータホップの数は低減しない。IPスイッチング法およびタグスイッチング法の両方において、各中間ホップはルータであるが、ただし、この中間ルータは、ホップカウンタの増分や発呼者IPアドレスの変更などの層3見出しについての適当な修正は行なわない。さらに、層3処理が幾つかのルータにおいて省かれるが、ただし、タグの置換（スワッピング）のために、各ルータは、IPパケットの見出しをパージング（構文解析）することでタグを識別し、パケット見出し部分に、置換すべきタグを書き込むことを要求される。

【0014】上述のモデルに対する代替として、IPスイッチング法の様々なバリエーションを、NHRP（次ホップルーティングプロトコル）との関連で用いることが可能であり、このようなシステムが、本発明の発明者らによって、1996年10月8日付けで出願された、本発明と譲受人を同一とする保属中の、“Internet NCP-Over ATM”という名称の米国特許出願第08/728,276号において開示されているので、これについても参照されたい。

【0015】上述のどのシステムも、IPパケットを網を横断して伝送するとき、IPパケットによって経験（経由）されるルータホップの数を低減し、同時に、各ホップにおける層3処理の量を低減することには失敗している。将来ますます多くのルータおよびウェブ（Web）サーバが展開されることが見込まれる中であって、網の成長を許し、データ伝送の性能を向上させるために、以前よりも増して高速なインターネットバックボーンが必要とされている。

【0016】

【課題を解決するための手段】本発明は、上述の網スキームおよびプロセスの短所および欠点を克服し、上述の課題を解決することを目的とする。この目的を達成するために、本発明は、IPトラヒックを伝送するための、IPパケットを網を通じて伝送した際に、IPパケット

によって遭遇されるルータホップの数を低減し、同時に、IPパケットが網を横断して伝送される際に遂行される層3処理の量を低減することが可能なインターネット中継システムおよび方法を開示する。

【0017】本発明の一つの実施例においては、中継網は、ソース網と通信し、ソース網からIPパケットを受信する入ルータを含む。この入ルータは、各IPパケットに、グローバル的に一意なラベルを付加する。このラベルは、IPパケットを網を横断して転送するのに用いられる。中継スイッチ網は、入ルータと通信し、入ルータからIPパケットを、ラベル内に含まれる情報に基づいて受信する。中継スイッチ網は、IPパケットをその伝送路に沿って、IPパケットに付加されたラベル内に含まれる宛先情報に基づいて転送する。出ルータは、中継スイッチ網からIPパケットを受信し、IPパケットを宛先網に転送する。宛先網は、いったんIPパケットを受信すると、IPパケットを宛先に転送する。

【0018】本発明のもう一つの実施例においては、中継網は、ソース網と通信し、ソース網からIPデータグラムを受信する入ルータを含む。この入ルータは、データグラムに、データグラムの意図する宛先に対応する情報を含むラベルを付加する。データグラムは、入ルータからスイッチ網に転送される。スイッチ網は、データグラムを出ルータに、出ルータが、あたかも、入ルータから1ルータホップしか離れてないかのように転送する。

【0019】

【発明の実施の形態】図3は、本発明の原理によるIP中継網の機能を簡略的に示す。IP中継網は、ローカリエリア網（“LAN”）/IP網100~105、IP中継ルータ（“IPRR”）110~115、およびIP中継スイッチ（“IPRS”）120~123を相互接続する。IPRR110~115は、網の“エッジ”を形成し、IPRS120~123は、IPRR間の接続性を提供する網の“コア”を形成する。LAN100~105は、典型的には、互いに接近して位置し、互いに網を形成するように接続された一群のユーザを表す。図3は、特定の数のルータ、スイッチ、およびIP網/LANを持つ本発明によるIP中継網の特定な構成を示すが、ただし、IP中継網のサイズは、要素の数あるいはこれら要素間の相互接続によって制限されるものではないことに注意する。

【0020】図3に示すように、LAN100~105は、おのおのIPRR110~115と通信する。ただし、網の構成によっては、ある特定のLANが、一つあるいは複数のIPRRと通信するようにすることも可能である。各IPRRは、マルチホームドホストであって、IPパケットが自身にアドレスされている場合は、そのパケットを受信し、IPパケットが別のホストにア

ドレスされている場合は、そのパケットを中継網内の一つあるいは複数のIPRS120~123を介して別のルータに転送する。各IPRR110~115は、受信された各IPパケットについて層3処理を遂行する。層3処理は、一般的には、各IPパケットの見出し部分内に含まれるIPアドレスのパーズング（構文解析）と、読み出しを意味する。この見出し部分内のIPアドレスには、ソースIPアドレスと、宛先IPアドレスに関する情報が含まれる。各IPRR110~115は、宛先アドレスを調べ、この宛先アドレスを、自身の網のIPアドレスと比較する。宛先アドレスとソースアドレスが同一のLAN上にある場合は、IPパケットは、その宛先アドレスを持つホストに直接にルートされる。宛先アドレスが同一のLAN上に接続されていない場合は、IPRR110~115内のそのLANと通信する特定の

【0021】LANからIPパケットを受信し、網への伝送を行なうIPRRは、“入口（ingress）”IPRRと呼ばれる。入口IPRRは、特定のIPパケットのアドレス見出し部分に含まれる情報に基づいて、ネットワーク層（層3）の所で、“ラベル”と呼ばれる固定長の見出しを付加する。このラベルが、IPパケットが網上を宛先に向けて転送されてゆく際に一つあるいは複数のIPRS120~123によって読み出される。IPRSは、入口IPRRからIPパケットを受信すると、そのパケットを、宛先ホストと接続されているLANと通信する宛先、すなわち“出口（egress）”IPRRに向けて転送する。IPRSは、IPパケットを、パケットの意図される宛先によって、その伝送路上のもう一つのIPRSに、あるいは、出口ルータに転送する。IPRRは、物理的に複数のIPRSに接続することも考えられる。この場合は、ルーティングアルゴリズムによって、IPパケットの伝送路内の最後のIPRSが、出口IPRRへの最適経路上に位置するIPRSとして決定される。

【0022】入口IPRRによって各IPパケットに付加されるラベルの重要な点は、このラベルの幾つかの部分が、中継網の幾つかの要素あるいは一群の要素を識別することである。この概念は、電話番号と、3桁のエリアコードが、電話が国のどの部分に位置するかを指定し、3桁の交換局コードがそのエリアコード内の地域（リージョン）を示し、最後の4桁の数字が、その交換局内の特定な回線を指定するという点で類似する。現存の網プロトコルシステムでは、パケットルーティング情報はランダムな識別子から構成されるが、これとは異なり、本発明による中継網において用いられるラベルは、宛先が網内のどこに位置するかに対応する情報を持つ。例えば、ラベルの第一の部分は、IPパケットが、網内

のどの一つあるいは一群のIPRSに向けられているかを示す。ラベルのもう一つの部分は、IPパケットが、IPRSのどのポートに、換言すればどの出口IPRRに、送られるべきかを示す。

【0023】各IPRSは、データリンク、すなわち層2デバイスとして機能する。すなわち、各IPRSは、各IPパケットに付加されたラベルを読み出し、そのIPパケットの伝送路に沿っての次のストップを決定する。ここで、このストップは、もう一つのIPRSでも、出口ルータでもあり得る。IPパケットが最終的に出口ルータによって受信されると、ラベルが除去され、IPパケットは、その宛先に転送される。タグスイッチング法やIPスイッチング法とは異なり、本発明によるIPRSは、IPパケットに付加されたラベルを、パケットをその伝送路に沿って転送する際に修正あるいはスワップ（置換）することはない。従って、本発明によるIPRSの転送動作は、ハードウェアによる高速処理が可能であり、このため、インターネットおよびインターネットアプリケーションによって要求される大きな転送速度およびスループットに答えることが可能である。

【0024】上述の概説は、以下の例によってより良く理解できるものである。説明の前に、以下の例は、単に、解説を目的とし、本発明の範囲および精神を制限することを意図するものではないことに注意する。ここでは、LAN100に接続された、例えば、パーソナルコンピュータであり得るソースホスト90、IPパケットを、LAN103に接続された、これもパーソナルコンピュータであり得る宛先ホスト95に送るものと仮定する。このIPパケットは、ホスト90から、従来の方法にて、LAN100を経て、IPRR110に送られる。IPRR110は、入口IPRRと見做され、IPパケット内のアドレス見出し部分を読み出し、宛先ホストが、IPRR110がパケットを直接に配送することが可能なLANには接続されていないことを知る。IPRR110は、IPパケットについて層3処理を遂行し、同時に、IPRR110のメモリデバイスに格納されているルーティングテーブル内に含まれる情報に基づいて、IPパケットに、次のホップのルータが出口IPRRと見做されるIPRR113であることを示すラベルを付加する。こうして、入口IPRR110は、出口IPRR113を、あたかも1ホップしか離れていないように見做す。

【0025】次に、このIPパケットがIPRS120に転送される。IPRS120は、例えば、ATM網であり得る。この転送プロセスは、IPRR110のルーティングテーブル内に格納されているアドレス情報と、これもIPRR110の転送テーブル内に格納されているアドレス情報とをマッピング（照合）することによって遂行される。IPRS120は、付加されたラベルを読み出し、IPパケットを、IPRS120の転送テ

ブル内の出口IPRR113に対応する情報に基づいて、IPRS122に転送する。IPRS122も、このラベルを読み出し、IPパケットを、IPRS122の転送テーブル内に格納された出口IPRR113に対応する情報に基づいて、出口IPRR113に転送する。出口IPRR113は、IPパケットを受信し、層3処理を遂行することで、パケットをリアセンブルし、これを、LAN103を介して宛先ホスト95に転送する。

【0026】本発明によるIP中継網は、入口IPRRによってIPパケットに付加されるラベルとして、幾つかの異なるフォーマットを扱うことが可能である。一つのバージョンのラベルにおいては、出口IPRRに対応する識別情報(IPRR ID)が用いられる。別のバージョンにおいては、ラベルは、出口IPRRと通信する最後のIPRSに対応する識別情報(IPRS ID)と、出口IPRRが通信の際に用いるIPRSポートに対応する情報(PORT ID)を含む。さらに別のバージョンとして、ラベル内にFLOW IDを含め、同一の宛先に向けられた複数のIPパケットを区別することも可能である。これは、IPパケットの伝送路に沿う各IPRSが、異なるIPパケットに対して、異なる品質のサービスを提供することを可能にする。例えば、FLOW IDによって、パケットが受信されたことを示す受信確認メッセージをソースに送り返す必要性を示すことも可能である。加えて、このFLOW IDは、IPパケットトラヒックを、中継網内の複数の最低コストの経路を経由してルーティングすべきことを示

DESTINATION IP ADDRESS (宛先IPアドレス)	FLOW ID	LAST IPRR ID	LAST {IPRS ID, PORT ID}
--------------------------------------	---------	--------------	-------------------------

テーブル1

テーブル1において、DESTINATION IP ADDRESS (宛先IPアドレス)は、IPパケットの見出しアドレス部分を読み出すことによって得られる。FLOW IDは、上述のように、入口IPRRによってローカル的に網管理者によって決定される政策に基づいて指定される。LAST IPRR IDは、宛先ホストがLANを介してそれと通信する出口IPRRを識別する。LAST {IPRS ID, PORT ID}は、オプションのフィールドであり、LAST IPRS IDは、出口IPRRと通信するIPRSを識別し、PORT IDは、出口IPRRが最後のIPRSを介して接続される物理ポートを識別する。採用されるテーブルのバージョンによって、各IPRR内に格納されるルーティングテーブルは、特定のIPパケットの宛先アドレスを出口IPRRにマッピングすること(LAST IPRR ID)も、あるいはIPRSおよび出口IPRRと通信する物理ポートにマッピングすること(LAST IPRS ID, PORT ID)も考えられる。これらルーティングテーブル内に出口ルータの

すために用いることも可能である。

【0027】IPRR110~115による、どのようなラベルを、ある出口IPRRに向けられたあるIPパケットに付加すべきかについての学習の仕方は、標準のIPルーティングプロトコル、例えば、OSPF (Open Shortest Path First; 空いた最短経路を最初に用いるプロトコル) に対する拡張に基づく。IPRR110~115およびIPRS120~123は、通常のIPルーティング情報/リンク状態トポロジー情報を配布するのみでなく、採用されているラベルのバージョンに応じて、IPRS ID情報およびPORT ID情報、あるいはIPRR ID情報を網全体に配布する共通のルーティング/トポロジー交換プロトコルに参加する。このプロトコルは、各IPRR110~115が自身のルーティングテーブルおよび転送テーブルを作成し、これを一つのメモリデバイス内に、あるいは各IPRR内の別のメモリデバイス内に格納することを可能にする。

【0028】ある与えられた宛先のIPアドレスに対して、各IPRR110~115内に格納されるルーティングテーブルは、IPデータグラムに付加されるべきラベルを提供する。このルーティングテーブルは、従来の網システムにおいては単に次のホップに当たるルータのアドレスのみを格納するのとは対象的に、各IPパケットをアドレス先に送り届けるための出口IPRRに対応する識別情報を含む。各IPRR110~115内に格納されるこのルーティングテーブルは、テーブル1に示すような一般構造を持つ。

【表1】

識別を加えることで各IPRRはルーティングされた網に関する完全な知識を持つこととなる。中継網に追加のIPRRが加えられる度に、各IPRR内のルーティングテーブルは、ルーティングアルゴリズムを用いて動的に更新される。各IPRR内に格納されるルーティングテーブルは、従来の技術の次のホップに当たるルータとは対象的に各層3宛先に向う途中の出口IPRRを含む。

【0029】加えて、本発明による中継網内に、IPRR110~115の一つを省略時のIPRRとして指定することにより、省略時のIPRRを実現することも可能である。IP中継網が、非-IP網に接続されている場合は、省略時のIPRRが、その非-IP網に向けられたトラヒックに対する出口IPRRと見做される。この省略時のIPRRは、IPトラヒックについて必要な層3処理を遂行し、IPトラヒックを非-IP中継網の最初のルータに転送する。例えば、パケットが、IP網

105から非-IP網155に送られた場合は、IPRR115が省略時のIPRRと見做される。IPRR115は、このパケットについて必要な層3処理を遂行し、次に、このパケットを非IPルータ150に転送する。非IPルータ150は、このパケットを、この

EGRESS IPRR ID	LAST IPRS ID	PORT ID(1)	PORT ID(2)
----------------	--------------	------------	------------

テーブル2

テーブル2において、EGRESS IPRR IDは、IPパケットが転送されるべき出口IPRRを識別する。LAST IPRS IDは、出口IPRRに物理的に接続されたIPRSを識別する。PORT ID フィールドは、特定のIPRRおよびIPパケットの宛先に依存するローカル的な意味を持つ。PORT ID(1)およびPORT ID(2)のフィールドは、IPパケットが送られる入口IPRRからのポートを識別する。換言すれば、PORT ID(1)およびPORT ID(2)のフィールドは、IPパケットが入口IPRRからどのIPRSに転送されるべきかを示す。これら2つの異なるPORT ID、つまり、PORT ID(1)およびPORT ID(2)フィールドは、ある一つのIPパケットの宛先に対して、複数の代替の（最低コストの）経路が存在し得るために用いられる。例えば、IPRR110の転送テーブルは、IPRS120に対応するPORT ID(1)と、IPRS121に対応するPORT ID(2)を持つ。パケットの出口IPRRがIPRR112である場合、入口IPRR110は、パケットを、IPRS120あるいはIPRS121のいずれかに転送する。どちらに転送するかは、現在網上を送信されており、この特定のIPパケットを出口IPRR112に接続する最低コストの経路に対して影響を与えるIPトラヒックの状況によって決定される。IPRRはIPパケットを出口IPRRにルートするために利用できる幾つかの代替ポートの一つを自身のローカルアルゴリズムに基づいて選択する。

【0031】各IPRS120～123は、メモリ内に、上述のテーブル2に示すのと実質的に同一の構造を持つ転送テーブルを格納する。PORT ID (PORT ID(1)およびPORT ID(2)) は、入口IPRRがIPデータグラムを出口IPRRと通信するIPRSを介して出口IPRRに向けてルートするために用いるべきポートを示す。入口IPRRが幾つかの異なる経路を利用できるのに類似し、PORT ID(1)およびPORT ID(2)は、IPパケットの伝送路内の次のIPRSへの代替の最低コストの経路を示す。IPRSがIPパケットを受信すると、IPRSは、ラベルのIPRS IDの部分を読み出し、これを自身の転送テーブル内のLAST IPRS IDの項目にマッピング（と照合）することで、IPRSの出方向のPORT IDを決定する。この出ポートは、パケットの宛先によって、パケットの伝送路内の次のIPRSに接続されていることも、あるいは、出口IPRRに接続さ

非-IP網155内の宛先に転送する。

【0030】各IPRR110～115内に格納される転送テーブルは、テーブル2に示されるような一般構造を持つ。

【表2】

れていることもある。こうして、PORT IDフィールドは、これがIPパケットを次のIPRSに、あるいは出口IPRRに、ルートするときに用いられるIPRSのポートを示すという意味でローカル的な重要性を持つ。さらに、PORT ID(2) が選択され、負荷バランシングが起動された場合は、FLOW IDフィールドのハッシングによってIPRSの特定の出ポートが選択される。FLOW IDフィールドは、さらに、例えば、FLOW ID(1)と命名される網を横断して機密な経路を要求するIPトラヒックと、FLOW ID(2)と命名される網を横断してあまり機密でない経路を要求するパケットとの間の区別を行なうために用いることも可能である。各IPRS内にルーティングテーブルを用いることで、IPパケットについて層3処理を遂行する必要性が回避される。

【0032】各IPRSおよびIPRR内の転送テーブルも網のトポロジーあるいはリンク状態に基づいて動的に更新される。こうして、網のトポロジーが、IPRRおよび/あるいはIPRSが追加あるいは削除されたことによって変化した場合、コア網を横断しての最適経路が変化し、転送テーブルもこれを反映して更新される。

【0033】図4は、本発明によるIP中継網内をIPデータグラムが伝送される際のIPデータグラムのソフトウェア処理の流れを示す。このプロセスは、ステップ200から開始され、ここで、IPRR110～115の中の入口IPRRと見做される一つが、LANあるいはIP網100～105に接続されたエンドステーションの一つからIPデータグラムを受信する。ステップ201において、入口IPRRがIPデータグラムのパージング（構文解析）を遂行し、特定のLAN100～105に接続された宛先のIPアドレスを読み出す。このIPアドレスは、IPデータグラムの見出し部分の中に含まれる。ブランチ205において、入口IPRRは、受信されたIPパケットの宛先IPアドレスに対応するルーティングテーブルの項目が存在するか否かを調べる。結果が肯定である場合は、ステップ207が遂行され、入口IPRRは、ルーティングテーブルの項目を検索し、IPパケットに付加されるべき“ラベル”を決定する。逆に、ステップ205における結果が否定であり、入口IPRR内に宛先アドレスに対するルーティングテーブルの項目が存在しないことが示された場合は、

17

ステップ209において、入口IPRRは、省略時のルータに対応するテーブルの項目を、出口IPRRとしてセットし、このプロセスは、上述のステップ207に進む。ステップ208において、IPRR100は、ラベルを層3処理を用いてIPデータグラムに付加する。条件分岐点211において、IPRR100は、データベース内に、LAST IPRR IDに対する転送テーブルの項目が存在するか否かを調べる。テストの結果が肯定である場合は、IPRR100は、条件分岐点271において、それが宛先ステーションへの最後のホップであるか否かを決定する。テスト結果が肯定である場合は、IPRR100は、ステップ291において、ラベルを除去し、IPデータグラムをその最終的な宛先に転送する。この場合、このプロセスは、ステップ250で退出する。

【0034】ステップ271におけるテスト結果が否定であり、受信エンティティが最後のIPRRではないことが示された場合は、IPRR100は、ステップ221において、IPデータグラムを転送テーブルから検索されたPORT IDによって識別されるポートに送る。ステップ211における結果が否定であり、出口IPRRに対する転送テーブルの項目が存在しないことが示された場合は、このプロセスは、ステップ250において退出する。

【0035】ステップ251において、入口IPRRからパケットを受信したIPRS120~123の一つは、自身のメモリ内に格納された転送テーブル内に転送

18

項目が存在するか否かを決定する。転送テーブルの項目が存在する場合は、IPデータグラムが、出口IPRRに向う次のIPRSに転送される。このプロセスは、ステップ211に戻り、ここで、伝送路上の各IPRSが、次のIPRSに対する転送テーブルの項目が存在するか否かを決定する。このプロセスが、経路に沿っての各IPRSにおいてデータグラムの配達完了するまで反復される。

#### 【図面の簡単な説明】

【図1】IPをATM網を通じて伝送するための従来の方法のブロック図を示す。

【図2】近道ルーティングを持つIP網のブロック図を示す図である。

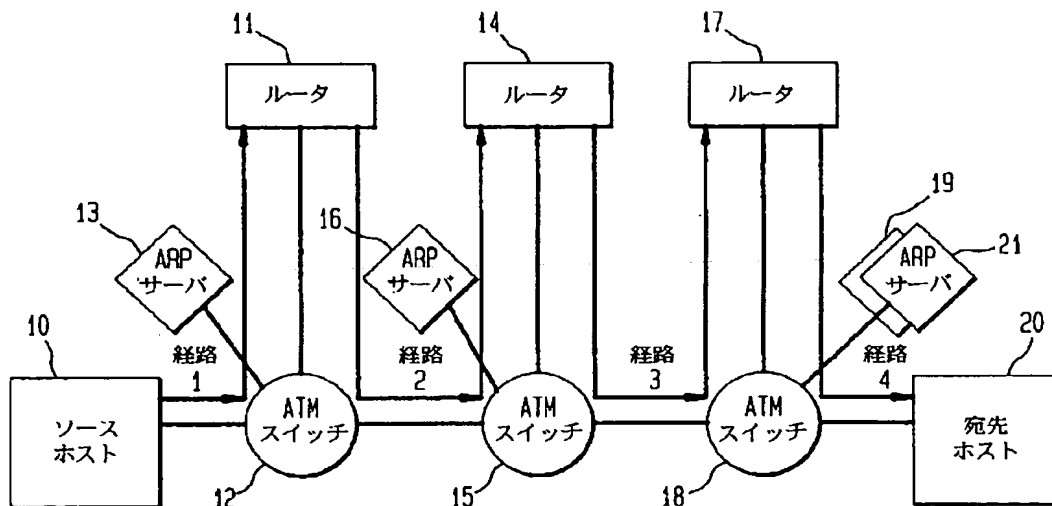
【図3】複数のルータが網のエッジを形成し、複数のスイッチが網のコアを形成する本発明による中継網の略図を示す図である。

【図4】図3に示される中継網を横断するときのIPデータグラムの処理の流れを示す図である。

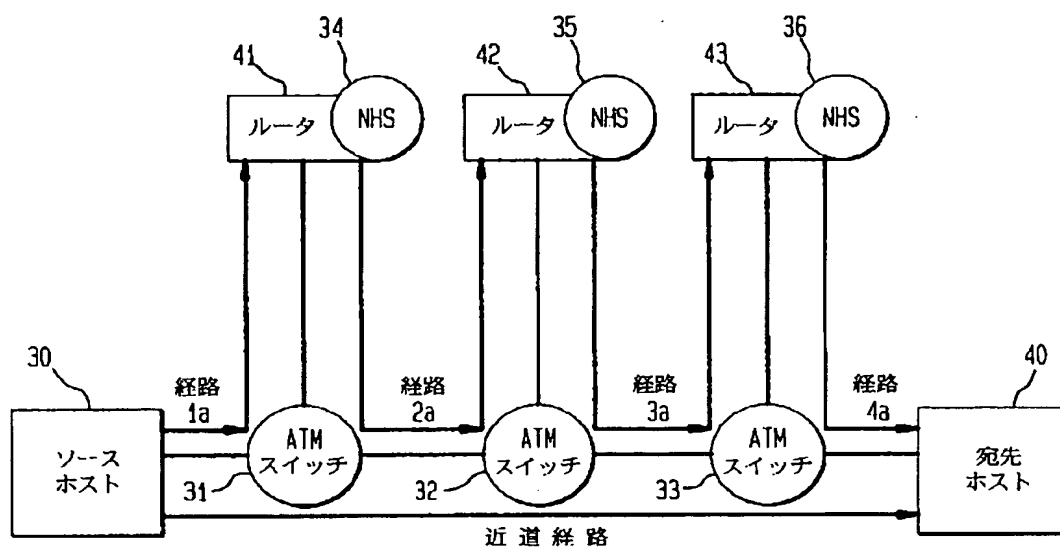
#### 【符号の説明】

90 ソースホスト  
95 宛先ホスト  
100~105 LAN/IP網  
110~115 IPRR (IP中継ルータ)  
120~123 IPRS (IP中継スイッチ)  
150 非IPルータ  
155 非IP網

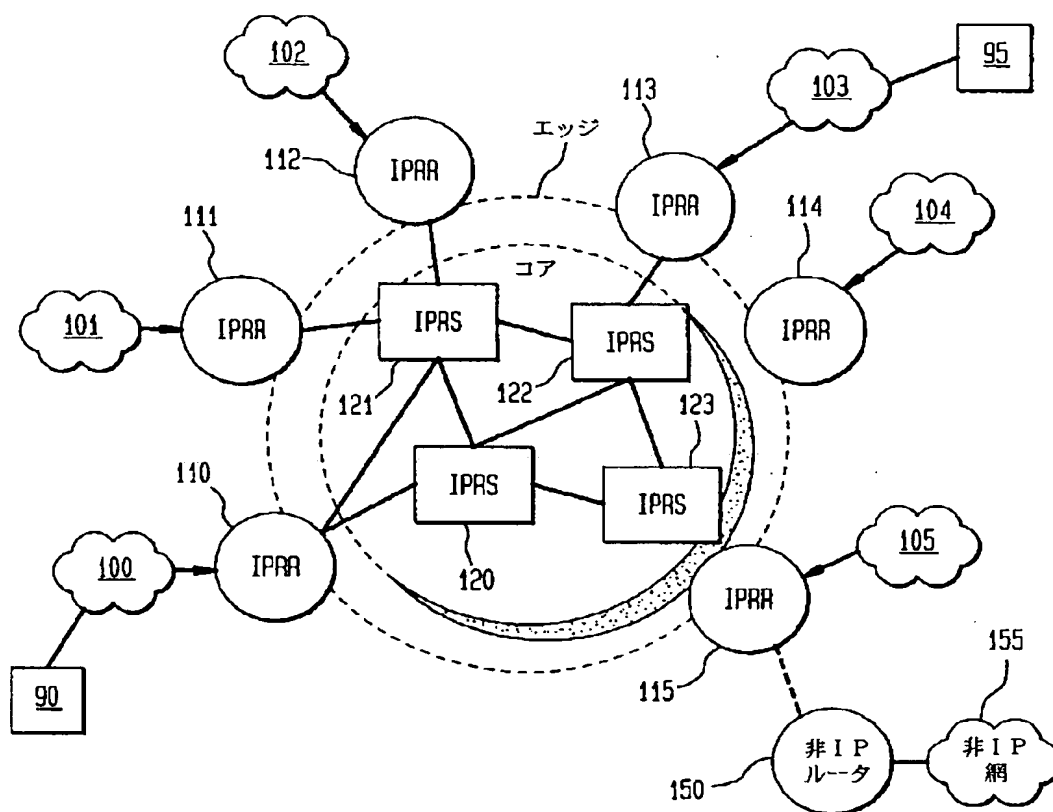
【図1】



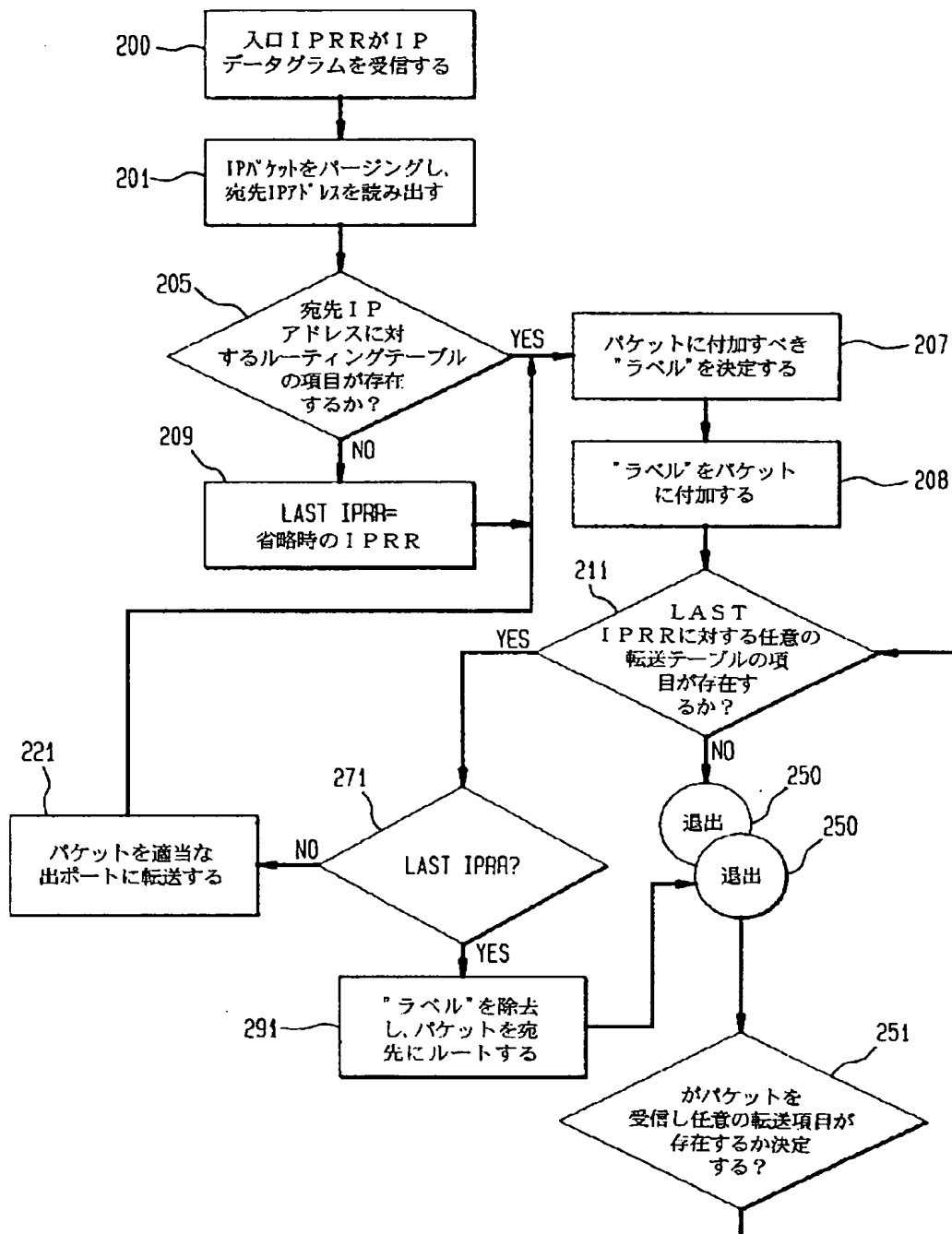
【図2】



【図3】



【図4】



フロントページの続き

(72) 発明者 ヴィクラム アール. サクセナ  
 アメリカ合衆国 07728 ニュージャージー  
 イ, フリーホールド, ベニントン ロード  
 281